LBL-14758 Rev.
Preprint

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

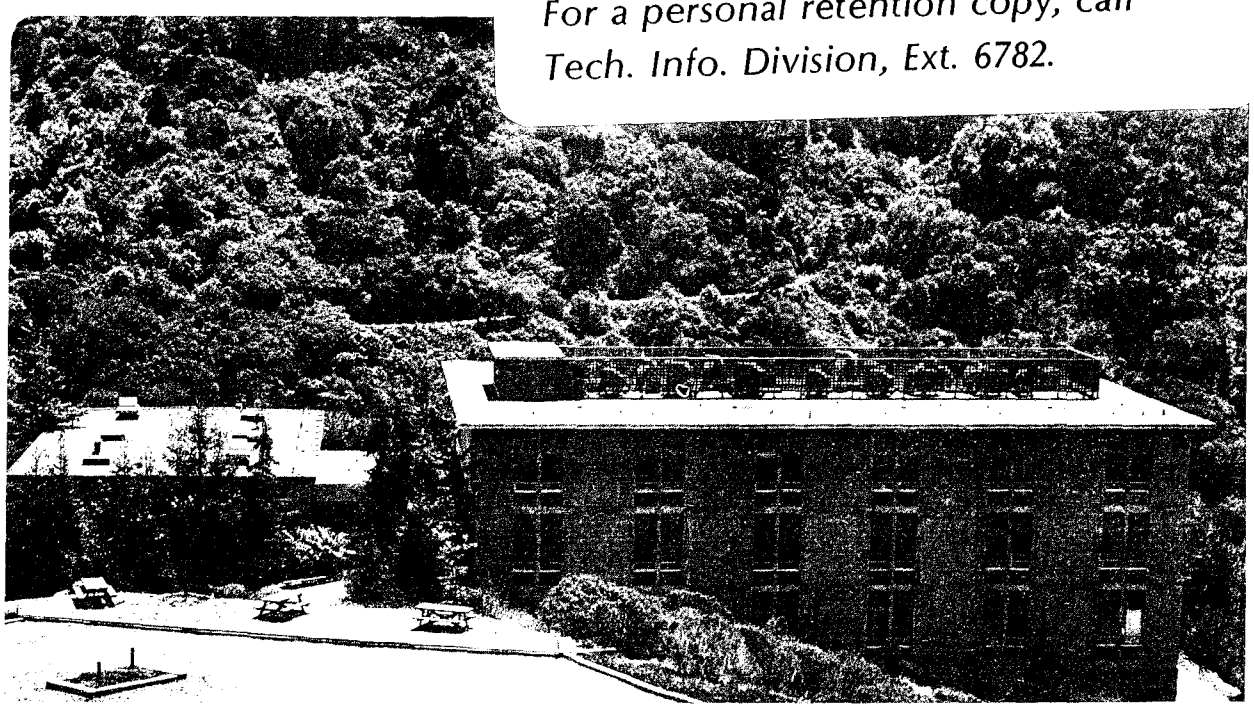
Materials & Molecular Research Division

To be submitted for publication

FOURIER ANALYSIS OF EXTENDED FINE STRUCTURE
WITH AUTOREGRESSIVE PREDICTION

J.Barton and D.A. Shirley

January 1985



Prepared for the U.S. Department of Energy under Contract DE-AC03-76SF00098

LBL-14758 Rev.

FOURIER ANALYSIS OF EXTENDED FINE STRUCTURE WITH AUTOREGRESSIVE PREDICTION

John Barton and D. A. Shirley

Materials and Molecular Research Division
Lawrence Berkeley Laboratory
and
Department of Chemistry
University of California
Berkeley, California 94720

FOURIER ANALYSIS OF EXTENDED FINE STRUCTURE WITH AUTOREGRESSIVE PREDICTION

John Barton and D. A. Shirley

Materials and Molecular Research Division
Lawrence Berkeley Laboratory
and
Department of Chemistry
University of California
Berkeley, California 94720

ABSTRACT

Autoregressive prediction is adapted to double the resolution of Angle-Resolved Photoemission Extended Fine Structure (ARPEFS) Fourier transforms. Even with the optimal taper (weighting function), the commonly used taper-and-transform Fourier method has limited resolution: it assumes the signal is zero beyond the limits of the measurement. By seeking the Fourier spectrum of an infinite extent oscillation consistent with the measurements but otherwise having maximum entropy, the errors caused by finite data range can be reduced. Our procedure developed to implement this concept applies autoregressive prediction to extrapolate the signal to an extent controlled by a taper width. Difficulties encountered when processing actual ARPEFS data are discussed. A key feature of this approach is the ability to convert improved measurements (signal-to-noise or point density) into improved Fourier resolution.

I. INTRODUCTION

Fourier transformation is a basic tool for spectroscopic data analysis in several contexts. Typically, Fourier transformation is used for harmonic analysis.[1-3] The spectroscopic measurement records an intensity while scanning energy; the Fourier transformation converts this energy spectrum into a frequency spectrum, reporting the amplitude and phase of a series of fixed frequency sinusoids which sum to the experimental result. If the physically significant part of the measurement has a distinctive frequency dependence, the signal frequencies can be isolated from irrelevant background or noise frequencies. Synthesis of the signal frequencies then yields a new energy spectrum whose interpretation may be simpler. For example, Extended X-ray Absorption Fine Structure (EXAFS) data are usually analyzed in this manner.[4,5]

Conceptually, Fourier analysis yields the amplitude and phase of each individual sine wave in a series which sums to give the spectroscopic signal. Of course, sine waves continue indefinitely while spectroscopic signals typically have a limited range. If the data analysis is restricted to a Fourier transform, this mismatch inevitably leads to a broadened Fourier spectrum: wide peaks appear for dominant frequencies, but adjacent peaks may overlap and the desired separation in frequencies may not be realized. With the Fourier methods currently used in spectroscopy[4-6] this finite-data-range broadening cannot be reduced by more careful measurements within a fixed interval. Thus if the measurement range is physically

restricted, then the ability of simple Fourier analysis to separate dominant frequencies will be limited.

Because of this broadening, the advantage of the explicit harmonic content analysis provided by a single Fourier transformation is offset by its lowering of frequency resolution. This broadening effect is extrinsic to the data set: it is inflicted on the data by forcing a clumsy method of analysis, because we force infinite sine waves functions to reproduce a finite length data sequence.[7,8] An implicit method for extracting the harmonic content (e.g., least-squares fitting the data) would provide the required frequency resolution. It is, moreover, also possible to realize the advantage of both approaches; viz, high frequency resolution and explicit analysis, by combining regression methods and Fourier analysis. Such an approach, for a particular spectroscopic method, is the subject of this paper.

To directly analyze Angle-Resolved Photoemission Extended Fine Structure (ARPEFS), a photoelectron diffraction phenomenon useful for surface structure determination,[9] we have found the frequency resolving power of the usual spectroscopic Fourier analysis to be inadequate, because the data range is limited. Fortunately, we have been able to adapt one of the new approaches to the Fourier analysis of physical measurements that allows higher Fourier resolution and can trade measurement precision for Fourier resolution. We shall report and discuss an adaptation of autoregressive prediction, also known as maximum entropy spectral analysis, which improves the Fourier

resolution by a factor of two in practical cases. Autoregressive prediction is widely used to process geophysical and acoustical measurements[1,3,10] when estimates of power spectra are required, but only short data sequences are available. We will demonstrate that autoregressive prediction can be used to extend the effective range of sinusoidal ARPEFS signals by an amount which increases with the signal-to-noise ratio. Although we apply this method to the analysis of ARPEFS, the method is directly applicable to EXAFS data or to other spectroscopies requiring high resolution Fourier transformations.

After ARPEFS is described in Section II, the taper-and-transform method of Fourier analysis is discussed in Section III. Autoregressive prediction is introduced in Section IV. The results are discussed in Section V, and a summary appears in Section VI.

II. ARPEFS

We shall demonstrate the autoregressive Fourier technique by applying it to ARPEFS data. In this section we briefly describe the essential physics of ARPEFS and discuss why high resolution Fourier analysis is required.

Angle-resolved photoemission extended fine structure is the oscillatory part of the photoemission current as a function of photoelectron kinetic energy.[9] Photo-excitation of an adsorbate core level gives an atomic-like (direct) outgoing photoelectron wave. Direct propagation of this wave to our detector would give an overall

atomic character to the differential cross section. Elastic scattering of this wave from substrate atoms leads to a new set of waves which can reach the detector and which interfere with the direct wave. For electron kinetic energies from about 50 to 500 eV, two conditions are met: single elastic scattering from ion cores dominates and the electron de Broglie wavelength corresponds to atomic dimensions. Thus, the interference modulation with kinetic energy can be used to derive the scattering path length and hence the position of the adsorbate atoms relative to the substrate.

The ARPEFS modulations are strongly dependent on the scattering angle, α_j, the angle between the photon polarization vector and scatterer, β_j, and the angle between the detector and the polarization vector, γ. In the simplest theory,[11] the modulations, $\chi(k)$, expressed as a function of the electron de Broglie wavenumber, k, are

$$\chi(k) = \sum_j A_j \cos[kr_j(1-\cos\alpha_j) + \phi_j], \qquad (1)$$

where

$$A_j = \frac{\cos\beta_j}{\cos\gamma} \frac{|f(\alpha_j)|}{r_j} e^{-L_j/\lambda_e k - \sigma_j^2 k^2(1-\cos\alpha_j)}$$

for 1s photoabsorption. If we call the polarization vector $\vec{\varepsilon}$, the emission vector $\vec{k}$, and the vector from the emitter atom to the j^{th} scatterer $\vec{r}_j$, then the parameters in this formula are:

β_j : angle between $\vec{\varepsilon}$ and $\vec{r}_j$

γ : angle between $\vec{\varepsilon}$ and $\vec{k}$

α_j : angle between $\vec{k}$ and $\vec{r}_j$

$|f(\alpha_j)|$: scattering amplitude for ion core j at $\vec{r}_j$

ϕ_j : scattering phase shift for ion core j at $\vec{r}_j$

λ : inelastic scattering length coefficient

L_j : total electron path in solid

σ_j^2 : mean square difference in displacement between emitter and scatterer j.

The argument of the cosine contains the geometrical information, $r_j - r_j \cos \alpha_j$. If the contribution from a single scatterer can be isolated, the scattering phase shift, ϕ_j, can be removed and the structure can be determined.

Because the single scattering theory is not valid for low wavenumber measurements and because the Debye-Waller factor, $\exp(-\sigma^2 k^2(1-\cos \alpha_j))$, reduces the intensity of the oscillations for high wavenumbers, the useful ARPEFS data range typically lies between 3 $Å^{-1}$-rad. and 12 $Å^{-1}$-rad. As we show in the next section this range may not be sufficient to resolve the nearest neighbor path lengths when normal Fourier analysis is applied.

III. THE TAPER-AND-TRANSFORM METHOD

To demonstrate our Fourier method we analyzed a harmonic sum

(Fig. 1a) made up of test data consisting of two sine waves with frequencies of 5 Å and 6 Å sampled 128 times in the interval from 4 to 11 $Å^{-1}$-rad. We added pseudo-random numbers to give a signal/noise ratio[12] of 2.8. Two important differences between this signal and our ARPEFS data--the k dependences of the amplitude A_j and of the phase ϕ_j--will be examined in Section V.

Direct application of the discrete Fourier transform,

$$g(p) = \frac{1}{N} \sum_{q=0}^{N-1} G(\tfrac{q}{N})\ e^{-i2\pi pq/N}, \qquad (2)$$

to the test sequence of N points [G] gives, via the Fast Fourier Transform,[2] a sparsely digitized Fourier spectrum, [g], shown in Figure 1b. The density of points in the Fourier spectrum can be increased by simply appending zeros to the sequence, [G], as Figures 1c and 1d illustrate, but ringing sidelobes--Gibbs oscillations--then appear, as a consequence of the finite length of the data sequence.[1,2] These oscillations obscure or confuse features in the experimental Fourier spectrum. They arise from the sharp truncation of the signal at the ends of the range. If y(p) is the sinusoid that we would get if we could measure an infinite range of data, then our experiment gives

$$b(p) = w(p)*y(p) \qquad -\infty < p < +\infty \qquad (3)$$

The box function, w(p), truncates the signal at the extremes of the

measurement interval:

$$\begin{aligned} w(p)&=0 && p < 1 \\ w(p)&=1 && 1 \leq p \leq N \\ w(p)&=0 && p > N \end{aligned} \qquad (4)$$

for N measurements. The Fourier transform of b is the convolution of the transform of the sine waves (delta functions) and the transform of the box ($\sin^2 x/x^2$). The sidelobes oscillations of the box transform are then superimposed upon the delta functions.

The usual approach for reducing these oscillations is termed "taper-and-transform" spectral analysis.[1] The sharp-edged box is replaced by a smooth weighting function whose Fourier transform does not contain large oscillations. This weighting function will broaden the Fourier spectrum as it reduces the sidelobe oscillations. Harris[13] and Nuttall[14] surveyed a variety of weighting functions and compared their performance by several criteria. For our purposes, the appropriate weighting function should have the highest possible resolution for a sidelobe-to-mainlobe ratio below the flat Fourier spectrum of the noise (assuming approximately normal distributed noise). Sidelobes falling below this level will have no more impact than the noise from the measurement. For sine waves, the Fourier signal-to-background will be the square root of half the number of data channels times the signal-to-noise ratio; this may be used as a rough guide for the weighting function selection.

As a measure of resolution we select the full width at half maximum valve and label it Δr. The width of the measurement, Δk, can

be related to this resolution as

$$\Delta r = \frac{2\pi}{\Delta k} f \qquad (5)$$

where the factor f depends on the weighting function. For a square window (no weighting) f = 1.21, and the sidelobe is .22 times the mainlobe. Harris gives f as the "6 dB resolution" and reports the sidelobe ratio in dB (20 times the $\log_{10}$ of the sidelobe ratio). Several of Harris' results[13] are collected in Table I and displayed in Figure 2; since Harris concentrated on weighting functions with very low sidelobes, we have extended his calculations to include weighting functions with sidelobes ~ 10 percent of the main lobe.

The weighting functions in Figure 2 fall in three groups. First, functions (a,b) which are flat in the center and fall smoothly to zero at the edges have the poorest resolution for a given sidelobe ratio. The shape of the roll off--Gaussian or cosine--seems to have little effect. Second, several functions (c,d,e) without variable parameters can be found which have 1-10 percent sidelobes but better resolution than the first group. Finally, the third set includes functions (f,g,h) which are theoretically optimal for mainlobe width versus sidelobe ratio by different measures.[13,14] For sidelobe ratios in the .1 to .01 range these weighting functions are equivalent.

From this last set we select the more familiar Gaussian weights and choose the Gaussian function width equal to 5/8 times the data range. This gives f = 1.6 and sidelobes equal to 3 percent of the

mainline. Figures3(a) and 3(b) illustrate the taper-and-transform results for the sine wave test spectrum using this weighting. The sidelobes will double while the mainlobe only narrows by 10 percent if we choose a Gaussian width equal to 3/4 of the data range.

With the resolution relation, equation 5, we can look forward to difficulties with real measurements. With the longest ARPEFS measurement range reported to date[9], $\Delta k = 6.5$ $Å^{-1}$-rad., the path-length resolution will be $\Delta r = 1.55$ Å. Nearest neighbor scattering atoms in that study appeared at path lengths of 1.96 Å, 3.2 Å, and 4.46 Å--these peaks cannot be resolved with taper-and-transform Fourier analysis.

IV. THE AUTOREGRESSIVE PREDICTION METHOD

The taper-and-transform Fourier method produces a Fourier spectrum of our signal only after we taper the signal toward zero at the edges of the observation interval. Beyond the observation interval this method therefore arbitrarily (albeit implicitly) assigns zero as the signal value, contrary to any reasonable expectation based on the sequence measured. In fact, most arbitrary choices for the signal in this region could be characterized as "unreasonable". This is another way of saying that we do not want the Fourier transform of our measured signal; we want the Fourier transform of a signal of which we have only a short segment. Proper selection of a weignting function can minimize the problems of a short data range, but this

does not address the underlying problem.

The autoregressive (AR) prediction approach to Fourier analysis proceeds with different assumptions about the data analysis problem.[1,3,10] In the AR method we assume that the data in the (limited) data range represent a few observations of an autoregressive process. By least-squares fitting these data we determine the process parameters and solve for the Fourier spectrum of the process. Because the range of the AR process is not limited to the observation interval, much better resolution is possible.

In an autoregressive process each data value, x_p, can be expressed as a linear combination of previous values,

$$x_p = - \sum_{q=1}^{m} a_q x_{p-q}. \tag{6}$$

The number m is called the "order" of the process; the coefficients a_q constitute an autoregressive filter. In modeling a data sequence with an AR process, a set of coefficients a_q and an order m must be calculated which can "predict" all the members of the data sequence. With the order less than the number of data points, the forward predictions in equation (6) and the backward predictions,

$$x_{p-m} = - \sum_{q=1}^{m} a_q x_{p-m+q} \tag{7}$$

form an overdetermined set of equations for the AR coefficients. The

structure of this set of equations is unusual since the autoregressive process employs data values to construct other data values.[10]

For n data values, we define a 2n-2m length vector, $\vec{b}$, containing the values from the left hand side of equations 6 and 7. Similarly, m AR coefficients form a vector, $\vec{a}$, and a (2n-2m) by m matrix, $\underset{\sim}{X}$, is constructed from the staggered data values as indicated on the right hand side. Then the least-squares equation for a is

$$\underset{\sim}{X}\vec{a} = -\vec{b} \qquad (8)$$

(see also ref. 10, page 249). These equations may be solved by fast recursive methods,[15] but, for the signal-to-noise ratios encountered in the analysis of extended fine structure, we find that the Singular Value Decomposition (SVD) method[16] for solving these equations to be more useful. As described in Ref. 17, the Singular Value Decomposition of a matrix $\underset{\sim}{X}$ gives

$$\underset{\sim}{X}_{2(n-m)Xm} = \underset{\sim}{U}_{2(n-m)x2(n-m)}\underset{\sim}{S}_{2(n-m)xm}\underset{\sim}{V}^{T}_{mxm}. \qquad (9)$$

where $\underset{\sim}{U}$ and $\underset{\sim}{V}$ are orthogonal matrices and $\underset{\sim}{S}$ is a diagonal matrix whose entries may be ordered by size. The value of this approach to solving least-squares problems has been discussed in detail by Lawson and Hanson,[17] and the application of spectral estimation is discussed by Tufts and Kumaresan.[16] Essentially, the SVD concentrates the significant signal content in the equations for the largest singular values. Then, when the solution for the original least-squares problem is constructed, only the largest singular values are used and

the remainder--those associated with noise--are discarded. Thus the solution for the AR coefficients is written

$$\vec{a} = - \sum_{i=1}^{P} \frac{\vec{V}_i}{\sigma_i^2} (\underset{\sim}{U}^T \vec{b})_i \qquad (10)$$

where $\vec{V}_i$ is the i^{th} row of $\underset{\sim}{V}$ and σ_i^2 is the i^{th} singular value.

The appropriate number of singular values, p, may be selected by visual inspection. Since the SVD of random numbers will be random, the ordered singular values with fall with a constant slope unless they contain information (see Figure 8, as discussed below). With the highly overdetermined AR system of equations, we can always see a section of constant slope for high index singular values: when the singular values rise above this slope, they contain information.

An AR process of order m has a Fourier power spectrum proportional to

$$\frac{1}{\left| 1 + \sum_{q=1}^{m} a_q e^{-2\pi i q p/N} \right|^2} \qquad (11)$$

Thus, one route to high resolution Fourier analysis proceeds as follows:

1. set $a_0 = 1$,
2. construct the least-squares equations and apply the SVD,

3. select the number of singular values and solve for a_q, q=1, m with equation 10,
4. set $a_q = 0$ for $q = M+1, q_{max}$ where q_{max} is a large power of 2, e.g. $q_{max}=2048$,
5. fast Fourier transform the full sequence $[a_q]$, and
6. invert the square modulus of the transform.

While this approach has the greatest potential resolution, it is difficult to apply to real data. The resulting peaks are all very sharp, making it difficult to distinguish spurious from real hidden peaks. The peaks are strong functions of the order chosen and of the signal-to-noise ratio in the data. Furthermore, only the power spectrum is retrieved; the phase information is not available.

For these reasons we have adopted a more conservative approach, suggested by reference 1, which sacrifices some resolution in favor of greatly enhanced reliability and control. This procedure is:

1. set $a_0 = 1$,
2. construct the least-squares equations and apply the SVD,
3. select the number of singular values and solve for a_q, q=1, m with equation 10,
4. use equation 6 to extrapolate the data sequence forward,
5. use equation 7 to extrapolate backward,
6. multiply the resulting sequence by a weighting function,
7. add zero value channels until the total number of channels is a large power of 2, e.g. 2048, and

8. fast Fourier transform the long sequence.

The Fourier coefficients derived from this procedure can be further analyzed with the usual Hilbert back transformation.[3,4] We have usually chosen an order equal to one half the number of data points, and we can typically extrapolate for approximately as many data points forward and backward as we originally measured.

The inherent control of this procedure comes in the examination of the extrapolated sequence. At some point in the extrapolation the new values begin to increase rapidly in amplitude and/or noise content (Fig. 3c). By placing the edge of our taper window at these points the unstable part of the extrapolation is eliminated. Furthermore, the window weights the extrapolated points significantly less than the real data values, moderating the effect of the new values on the final spectrum. This last advantage is crucial for practical spectroscopic signals which are not exact sinusoids.

An example of the extrapolation is shown in Figure 3c, and its effect on the Fourier spectrum is shown in Figure 3d. The signal in Figure 1a was fitted to an AR process of order 64 and 4 singular values were required (as expected for 2 real sinusoids[16]). Extrapolation gives Figure 3c. Figure 3d dramatically illustrates the potential of this method for increasing resolution in Fourier analysis.

At this point it is useful to note that the ARP-Fourier transform method is <u>not</u> a "deconvolution" of the data which can produce spurious peaks through unreliable resolution enhancement. As we illustrated in

Figure 4, our net process solves a problem with the taper-and-transform Fourier method. In Figure 4a it is obvious on visual inspection that more than one frequency is present, but the Fourier transform will have Gibb's oscillations. When the taper (weighting function) is applied as in Fig. 3b, the beat structure is lost while the Gibb's oscillations in the Fourier transform are suppressed (Figure 3(b)). From this perspective the unadorned Fourier transform and the taper-and-transform process are clumsy operations that obscure the frequency information inherent in the data. When the ARP is applied, Figure 4(c), data on the ends of the measurement are no longer lost when the window function is applied, Figure 4(d).

Our overall procedure requires three parameters: the number of singular values, the number of AR coefficients, and the final taper width. As discussed above and in reference 17, the number of singular values may be determined by inspection. For poor signal-to-noise conditions, the size and variability of the singular values associated with noise will make this choice more difficult. Autoregressive orders between N/2 and 3N/4 are are recommended by Tufts and Kumaresan.[16] Our choice of a taper width at just less than twice the measured data width reduces the importance of our choice in the first two parameters.

Up to this point we have assumed that our measurement can be successfully approximated by an autoregressive process. In reexamining this point we divide the question in two parts: i) how closely can a cosinusoidal series be represented by an autoregressive

process, and ii) how closely does a cosinusoidal model fit ARPEFS data? For the first part we can note the discussion of Ulrych and Ooe[10]. Beginning with a finite difference equation for a sinusoidal series, they demonstrate that such a series can be represented by a combination autoregressive, moving average (ARMA) model; they also show that such an ARMA model can be represented by an infinite order pure autoregressive model. Numerical work by Tufts and Kumaresan[16] supports the conclusion that AR models can represent sinusoidal series; their method can give resolution near the theoretical limit even for low signal-to-noise ratios.

The second question is more difficult to address, but it impacts every method of harmonic analysis applied to ARPEFS. Specifically, if the cosine form breaks down, the taper-and-transform approach will fail as the autoregressive approach does. We will examine some of the possible problems in the next section.

V. DISCUSSION

Two important features neglected in the sine wave model spectrum are the amplitude and phase variation with k in real ARPEFS data. The sine wave model spectrum neglected any variation in frequency due to nonlinearity in ϕ_j and any variation in amplitude due to $|f(\alpha_j)|\exp[-\sigma^2k^2(1-\cos\alpha_j)-L_j/\lambda k]$.

To examine a model containing realistic amplitude and phase functions on a scale similar to our data we have generated a spectrum

by adding noise to

$$\chi(k) = \cos(173^\circ)|f(173^\circ)|\cos\,[4.46k + \phi(173^\circ)]\; e^{-.02k^2-5.3/k}$$
$$+\; 2\cos(116^\circ)|f(116^\circ)|\cos\,[3.21k + \phi(176^\circ)]\; e^{-.014k^2-5.3/k}, \qquad (12)$$

where f and ϕ are derived from summed partial-wave phase shifts.[18] Direct application of the AR prediction gives the result in Figure 5a and the Fourier transform in Figure 5b. The increase in amplitude at low k in the linear prediction is a consequence of the amplitude structure for scattering through 116°: $|f(116^\circ)|$ peaks at ~ 5 Å^{-1}-rad as shown in Figure 6. The AR method presumes that this is a rising signal and continues the trend to lower k. At higher k, the AR method tries to force this single decaying frequency to be modeled by infinite sine waves: it must sum two nearby frequencies to simulate the amplitude decline. The Fourier spectrum then contains a split peak for this scattering event.

The rising low k amplitude effect can be recognized in the predicted spectrum and remedied by analyzing $k\chi(k)$. The k weighting helps to cancel the decline of $|f(\alpha_j)|$ at higher k and has been used extensively for analysis of EXAFS data.[5] This weighting evens out the linear prediction shown in Figure 5c, and the resulting Fourier transform amplitudes (Fig. 5d) are more similar to the average amplitudes of the signals within the real measurement range.

Whatever weighting is employed, the important separation of the

Fourier frequencies is still effected by the autoregressive prediction followed by Fourier transformation. The amplitude variation places an upper limit on the resolution obtainable from the AR analysis of real data. When the amplitude function falls with the same shape as the beat envelop, then the AR analysis cannot distinguish between them.

Variation of the frequency with k violates the stationary assumption in the application of the autoregressive model. Thus ARPEFS peaks with phase functions strongly nonlinear in k will be modeled incorrectly, probably being represented as more linear than they really are. If the phase has an average slope at the beginning of the data range which is different from its average slope at the end, the extrapolation procedure sometimes yields a slightly doubled or asymmetric peak which must not be mistaken for two.

The frequency variation may also explain the empirical selection of a large process order m. In the usual application of the AR technique[10] the order is chosen by some criterion based on the prediction error; that is, the difference between the linear prediction and the data values. While this criterion can give a prediction filter for pure sinusoids in the presence of noise, valid for infinite range, we seek an adequate representation of a more complex oscillating signal over a small range. Our signal does not result from any autoregressive process, and a large order may model nuances of nonlinear phase and noise.

The impact of modelling this non-stationary signal with an

autoregressive filter is minimized because we do not rely on the Fourier spectrum itself for the final analysis. Following Martens,[4] we apply a Hilbert transformation[3] to our data. From the complex exponential form of the cosine

$$A_j \cos(p_j k + \phi_j) = \frac{A_j}{2} e^{i(p_j k+\phi_j)} + \frac{A_j}{2} e^{-i(p_j k+\phi_j)} \qquad (13)$$

we see that the transform of the cosine is real and peaked near p_j and $-p_j$. By using only the positive frequency components, a complex back transform gives

$$\frac{A_j}{2} e^{i(p_j k+\phi_j)} = \frac{A_j}{2} \cos(p_j k+\phi_j) + \frac{iA_j}{2} \sin(p_j k+\phi_j) \qquad (14)$$

The amplitude and phase functions of the original cosine wave can be derived as the amplitude and phase of this complex sequence. For our signal the actual cosine argument is $k(r_j-r_j) + \phi_j$, so we subtract the potential phase shift, ϕ_j, and fit the resulting sequence to a line. The slope of this line gives the averaged geometrical position we seek. The crucial point is this: we only use the cosine phase function in the region of k where we made actual measurements. Thus the entire AR prediction Fourier analysis serves only to isolate a single frequency. The position and amplitude of the Fourier peaks need not be accurate for us to obtain accurate geometries.

This complex backtransformation procedure does introduce one important source of error: we lose the wings of the Fourier peak

spectrum for the oscillation we isolate. This implies that the non-linear phase information and strong amplitude dependence of the ARPEFS oscillation will be missing in the backtransformed peak: we necessarily derive an averaged frequency and smoothed amplitude dependence. While not ideal, this result is certainly preferred to mixing the arguments of two different cosine oscillations.

Since there are a large number of variable parameters in even this simple model, we cannot yet give a complete analysis of the effects of background subtraction and signal/noise ratio. Generally, the AR prediction produces a "peakier" spectrum than one might imagine being correct.[1,10] Thus, errors in background subtraction appear as small peaks at harmless low r_j values. When the beat pattern of two peaks approaches the width of the actual measurement range, then errors in background subtraction may interfere with resolution.

Signal/noise ratios greater than two allow approximately double the resolution of the taper approach, with errors in geometry of < 0.02 Å. Errors increase rapidly for signal/noise ratios falling below 1. Until more experience is acquired with the AR method, prudence suggests examination of these effects for model spectra closely mimicking the actual data before assigning error limits.

As a practical example of the improved analysis of ARPEFS data, we have analyzed[9] the modulations (Figure 7a) in the sulfur 1s photoemission intensity emitted along the [110] direction from a c(2x2)S/Ni(100) adsorbate system. The Fourier transform via the taper approach shows distinct peaks (Figure 7b), but each peak is an average

of several path-length differences. The singular values for the application of an 128 order AR prediction are shown in Figure 8. The slope of the singular values is roughly constant—as indicated by the plotted derivative--above singular value 17. Thus 17 principal vectors were used to construct the AR filter. The AR prediction is shown in Figure 7c, and the Fourier transform gives Figure 7d. Now the individual peaks are clearly separated and they can be assigned to scattering path-length differences.[9]

VI. SUMMARY

Autoregressive prediction provides a method for greatly increasing the resolution of Fourier analysis of sinusoidal data. Using the extrapolate-taper-transform method described here, we can always do as well or better than the taper-transform approach. If the signal/noise ratio is so poor that the extrapolation fails immediately, then the AR procedure reverts to the usual taper method. For all other cases the resolution is improved. Furthermore, the method is easy to implement, computationally efficient, and controllable.

The resolution improvement afforded by the autoregressive prediction method scales with the quality of the experimental measurements. Low precision or widely spaced measurements do not contain enough information to accurately detemine the autoregressive coefficients. Our moderate precision measurements yield moderate precision autoregressive coefficients; our coefficients allow successful extra-

polation as we have demonstrated, but they are not precise enough for the analytic power spectrum formula.

Two improvements in the application of autoregressive prediction to spectroscopic data require further investigation. First, the statistical accuracy of the data values can vary significantly across a spectrum; the least-squares fit of the autoregressive coefficients should be weighted accordingly. Second, the autoregressive method assumes equal intervals between measurements; for ARPEFS we do not have equally spaced data. This problem is more difficult: the AR process given in equation (6) steps by a single fixed amount. However, there should be some AR process whose Fourier spectrum closely approximates the Fourier specrum of our data even if our measurements do not fall on an even mesh. Such questions are being examined in the signal processing literature, e.g., references 20 and 21, and new methods should be available soon.

Our final procedure is empirical for the same reasons the familiar taper-and-transform method is empirical. Ideal frequency analysis--the separation of our signal into each component oscillation--cannot be accomplished with noisy, finite-range measurements. Furthermore, harmonic analysis is only approximately valid for our spectroscopy: nonlinear phase shifts and energy-dependent scattering power preclude pure sine-wave signals. The procedure we have described here will, however, give a useful, high-resolution Fourier transform from real spectroscopic signals.

Formulation of the autoregressive prediction method from the

vantage of information theory has led to its description as maximum entropy spectral analysis.[7,8,10,22] Faced with the problem of estimating the Fourier transformation of an oscillatory signal given only a short measurement range, the autoregressive method fits a general oscillatory model to the measurements. The resulting over-determined set of equations are reduced by maximizing the entropy of the model. Thus, of all the possible models which give the same least-squares error, we select the model which adds the least new information, i.e. the one with the most signal entropy.

Data analysis methods can generally be compared by examining the information they add to the measurement. The AR method assumes that the data represent a process whose Fourier spectrum does not change outside the data sequence: it attempts to add no new information.[7,8,10] The taper-and-transform approach added the "information" that the signal was zero where it was not measured; this is contrary to any reasonable expectation. Directly fitting the data to a model of the physical process (eqn. (1)) would be the ultimate addition of information, but small uncertainties in the measurement and in the model usually prevent this approach[5] from being successful.

Note that extrapolation after direct physical model fitting has a different meaning than our AR prediction. Extrapolating by evaluating a physical model estimates a physical signal. The AR extrapolation does not estimate a physical signal; instead it reflects the frequency content over the original measured interval. The AR prediction

estimates an autoregressive model, not a physical one. We are not attempting to predict a measurable quantity; the extrapolation is merely one step in a harmonic analysis of our data.

Finally we note that this conservative approach to AR Fourier analysis can also be applied to a number of spectroscopic problems. Extended X-ray Absorption Fine Structure (EXAFS) has a nearly identical form to eq. (1), and the autoregressive prediction would allow high resolution Fourier analysis of more general utility than the beat method of Martens.[19] Many problems in spectroscopic deconvolution via the Fourier transform can also benefit from this AR approach. Direct AR power spectral analysis has been successfully applied to this problem,[23] but the danger of spurious peaks is particularly acute when we are seeking resolution enhancement. An extrapolation-taper procedure would allow a more controlled, albeit more moderate resolution enhancement.

Acknowledgements

The authors are grateful for instructive discussions with J. D. Klein and S. W. Robey.

This work was supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, Chemical Sciences Division of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098. It was performed at the Stanford Synchrotron Radiation Laboratory, which is supported by the Department of Energy, Office of Basic Energy Sciences and the National Science Foundation, Division of Materials Research.

References

1. S. M. Kay, S. L. Marple, Jr., Proc. IEEE 69, 1380 (1981).

2. E. O. Brigham, The Fast Fourier Transform, Prentice-Hall, Englewood Cliffs, N.J., 1974.

3. J. F. Claerbout, Fundamentals of Geophysical Data Processing, McGraw-Hill, New York, 1976.

4. G. Martens, P. Rabe, N. Schwentner, and A. Werner, Phys. Rev. B 17, 1481 (1978).

5. P. A. Lee, P. A. Citrin, P. Eisenberger, B. M. Kincaid, Rev. Mod. Phys. 53, 769 (1981).

6. Z. Hussain, D. A. Shirley, C. H. Li, S. Y. Tong, Proc. Natl. Acad. Sci. USA 78, 5293 (1981).

7. E. T. Jaynes, Proc. IEEE 70, 939 (1982).

8. J.A. Cadzow, Proc. IEEE 70, 907 (1982).

9. J. J. Barton, C. C. Bahr, Z. Hussain, S. W. Robey, J. G. Tobin, L. E. Klebanoff, and D. A. Shirley, Phys. Rev. Lett. 51, 272 (1983).

10. S. Haykin, ed. Nonlinear Methods of Spectral Analysis, Topics in Applied Physics, v. 34, Springer-Verlag, Berlin, 1983.

11. Derived by analogy with P. A. Lee and J. B. Pendry, Phys. Rev. B 11, 2795 (1975).

12. We define our signal-to-noise ratio as the ratio of the root mean square (rms) signal power divided by the rms noise power. For a sine wave with amplitude A, the rms signal level is 0.707A; the rms noise power equals the standard deviation of the noise, σ. The signal-to-noise ratio for a sine wave in noise is

then equal to $0.707A/\sigma$. The idealized Fourier spectrum for a sine wave in noise would be a single peak of height $A/2 + \sigma/\sqrt{N}$ with a flat background level equal to $\sigma/\sqrt{N}$.

13. F. J. Harris, Proc. IEEE 66, 51 (1978).

14. A. H. Nuttall, IEEE Trans. Acoust. Speech and Signal Proc., ASSP-29, 84 (1981).

15. L. Marple, IEEE Trans. Acoust. Speech, Signal Process. vol. ASSP-28, 441 (1980). This reference reports a FORTRAN program to calculate the AR coefficients.

16. D. W. Tufts and R. Kumaresan, Proc. IEEE, 70, 975 (1982).

17. C. L. Lawson, R. J. Hanson, Solving Least-Squares Problems, Prentice-Hall, Englewood Cliffs, NJ (1979). The computer programs published in this reference are available commercially from IMSL, NBC Building, 7500 Bellaire Blvd., Houston, Texas 77036.

18. C. S. Fadley, private communication.

19. G. Martens, P. Rabe, N. Schwentner, A. Werner, Phys. Rev. Lett. 39, 1411 (1977).

20. J. P. Schott and J. H. McClellan, IEEE Trans. on Acoust., Speech, Signal Processing, vol. ASSP 32, 410 (1984).

21. T. A. Ramstad, IEEE Trans. on Acoust., Speech, Signal Processing, vol. ASSP 32, 577 (1984).

22. A. Van den Bos, IEEE Trans. on Information Theory, IT-17, 493 (1971).

23. R. P. Vasquez, J. D. Klein, J. J. Barton, F. J. Grunthaner, J. Electron Spectrosc. Relat. Phenom. 23, 63 (1981).

Table I.

Resolution factors and sidelobe ratios for Fourier weighting functions. For a data range of Δk, the full width of the Fourier amplitude mainlobe for these weighting functions is ΔR where $\Delta r \Delta k = 2\pi f$. The ratio of the maximum sidelobe peak value to mainlobe peak is SL. These results are displayed in Figure 2.

Curve in Figure 2	Weighting Function	Formula $h = \Delta k/2$	B	f	SL
	None	$w(x) = 1$		1.20	.22
(a)	Tukey;[10] for B=0 Hanning	for $\lvert x-h \rvert < Bh$ $w(x) = 1$; for $x-h > Bh$ $w(x) =$ $\frac{1}{2} - \frac{1}{2}\cos\left[\pi \left(\frac{x-h-h}{Bh}\right)\right]$	.75 .66 .50 .33 .25 .00	1.38 1.43 1.57 1.72 1.80 2.00	.21 .20 .18 .13 .11 .03
(b)	Gaussian Step, or Error function	$\left[\frac{1}{2} + \frac{1}{2}\,\mathrm{erf}\left(\frac{x-Bh}{\sqrt{2}\,Bh}\right)\right]\left[\frac{1}{2} - \frac{1}{2}\,\mathrm{erf}\left(\frac{2h-Bh-x}{\sqrt{2}\,Bh}\right)\right]$	.125 .250 .333 .500 .750	1.37 1.47 1.72 2.02 2.18	.21 .18 .14 .04 .02
(c)	Riesz[10]	$1.0 - \left\lvert \frac{x-h}{h} \right\rvert^2$	-	1.59	.09
(d)	Cosine[10]	$\cos\left[\left(\frac{x-h}{h}\right) \pi/2\right]$	-	1.65	.07
(e)	Riemann	$\left[\sin\left(\frac{x-h}{h}\right) \pi\right]/\left(\frac{x-h}{h}\right) \pi$	-	1.74	.05

Table I continued.

Curve in Figure 2	Weighting Function	Formula $h = \Delta k/2$	B	f	SL
(f)	Van Der Maas	$\frac{B\, I_1[B\sqrt{1-((x-h/h)^2}]}{2h\sqrt{1-((x-h)/h)^2}}$	.5	1.14	.26
			1.0	1.17	.20
			2.0	1.28	.15
			3.0	1.38	.10
			3.5	1.43	.08
			4.0	1.51	.06
		$+\frac{1}{2}\,\delta(x-2h) + \frac{1}{2}\,\delta(x)$	5.0	1.65	.03
(g)	Gaussian		.80	2.22	.001
			1.00	1.82	.01
			1.24	1.58	.03
			1.50	1.45	.07
		$\exp\{-\ln 2[\frac{2(x-h)}{Bh}]^2\}$	2.0	1.33	.12
			2.4	1.29	.15
			3.0	1.26	.17
			4.0	1.23	.19
(h)	Kaiser-Bessel		.5	1.21	.21
			1.0	1.24	.18
			1.5	1.29	.15
		$\frac{I_0[B\sqrt{1-((x-h)/h)^2}]}{2h}$	2.0	1.36	.12
			3.0	1.50	.07
			3.5	1.58	.04
			4.0	1.65	.03
			5.0	1.80	.01

Figure Captions

Figure 1. (a) Sum of two sine waves, periods of 5 and 6 Å, plus 10 percent pseudo-Gaussian noise. (b) Fourier amplitude of the sequence in 1(a). (c) Extension of the sine waves of (1a) by appending zeros. Set above the signal is a plot of the weighting window function; it has a baseline of zero and a height of one. (d) Fourier amplitude of Fig. 1(c).

Figure 2. Resolution factor versus sidelobe-to-mainlobe ratio for several weighting functions. Abcissia is f in $\Delta r \Delta k = 2\pi f$; for a data range of 6.3 $Å^{-1}$-rad., f will be the Fourier resolution in Å. Ordinate is the maximum sidelobe peak value divided by the mainlobe peak. The plotted values are given in Table I. The point at f = 1.21 and sidelobe = .22 represents an unweighted Fourier transform. The weighting functions are given in Table I. (a) Tukey weighting, ref. 10, pg. 66. This function is flat in the center and rolls off as a cosine on the data extremes. (b) Gaussian Step or Error function. Similar to (a) but using a Gaussian roll-off. (c) Riesz polynomial, ref. 10, pg. 65. (d) Riemann weighting, ref. 10, pg. 65. (e) cosine weighting, ref. 10, pg. 60. (f) Van der Maas weighting, ref. 11, pg. 90. (g) Gaussian weighting, ref. 10, pg. 69, (h) Kaiser-Bessel weighting, ref. 11, pg. 89.

Figure 3. (a) Extended sine wave from Fig. 1(c) and, set above, the weights used for taper-and-transform Fourier analysis. The base of the weighting function is zero and its peak is one. (b) Fourier transform of sine wave times weights from Fig. 3(a). (c) Autoregressive prediction of the signal in Fig. 1(a), using an order m= 64, half of the 128 points. The new weights is set above. (d) Fourier amplitude of the product of the prediction results and weights from Fig. 3(c).

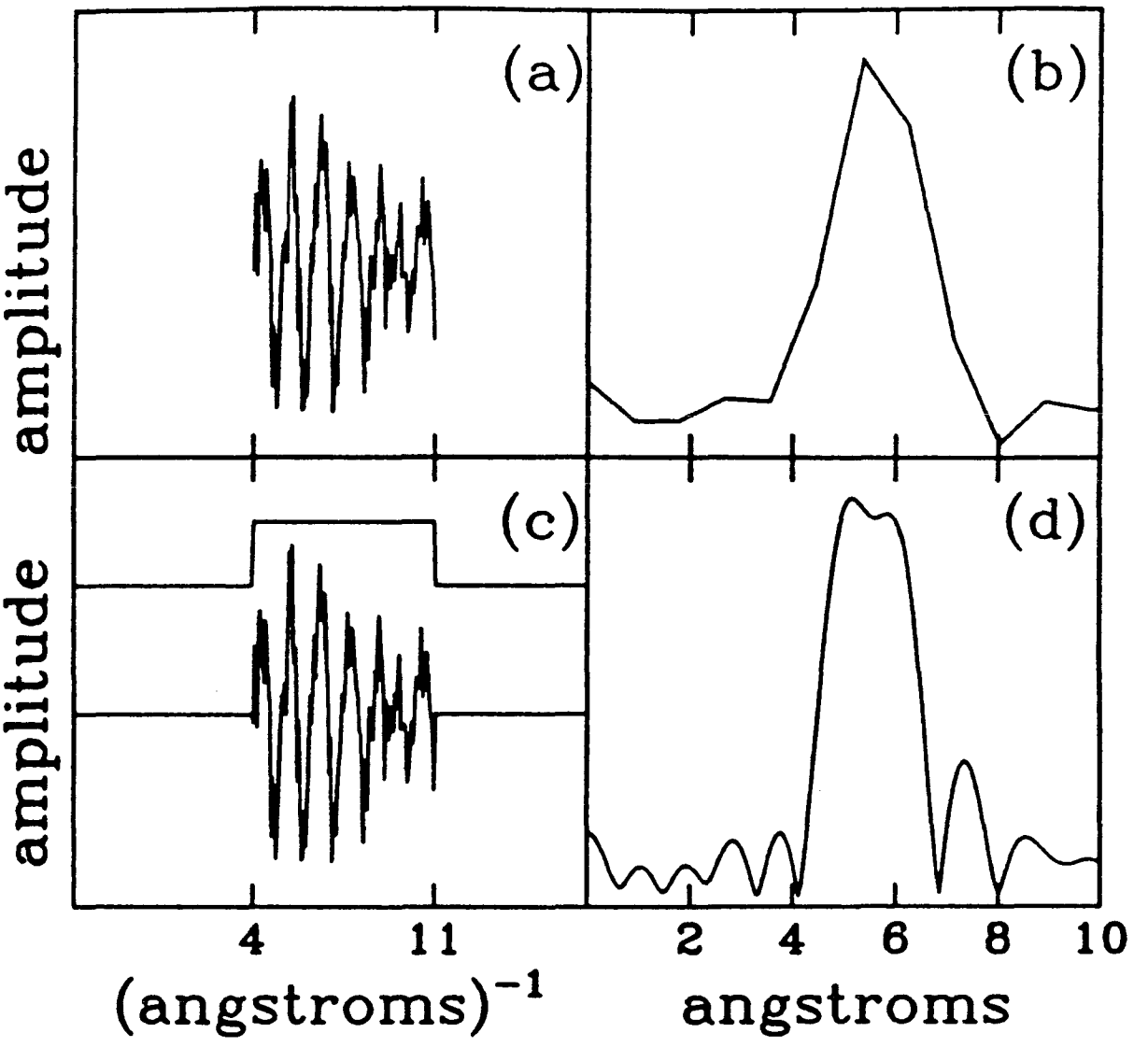
Figure 4. Weighting function interaction with autoregressive prediction. (a) Test sequence of two sine waves and noise as in Fig. 1. Note the beat structure. (b) Data from (a) times Gaussian weights. Gaussian width is 5/8 times the data range. Note the loss of beat structure. (c) AR prediction of the data in (a). (d) AR prediction from (c) times Gaussian weights. Gaussian width is 5/8 times the extended data range. Note the reduced emphasis of the extrapolated region.

Figure 5. (a) Autoregressive prediction of a simulated signal from equation (9). (b) Fourier amplitude of Fig. 3(a) times Gaussian weights. (c) Autoregressive prediction of k times the simulated signal in Fig. 3(a). (d) Fourier amplitude of Fig. 3(c) times Gaussian weights.

Figure 6. Magnitude of the scattering amplitude, $|f(\alpha,k)|$, for Ni atom at $\alpha = 116^\circ$ and $\alpha = 173^\circ$. The mild amplitude behavior of the scattering for 173° gives a simple Fourier peak shape; the steep drop at high k for scattering through 116° leads to a doubled Fourier peak.

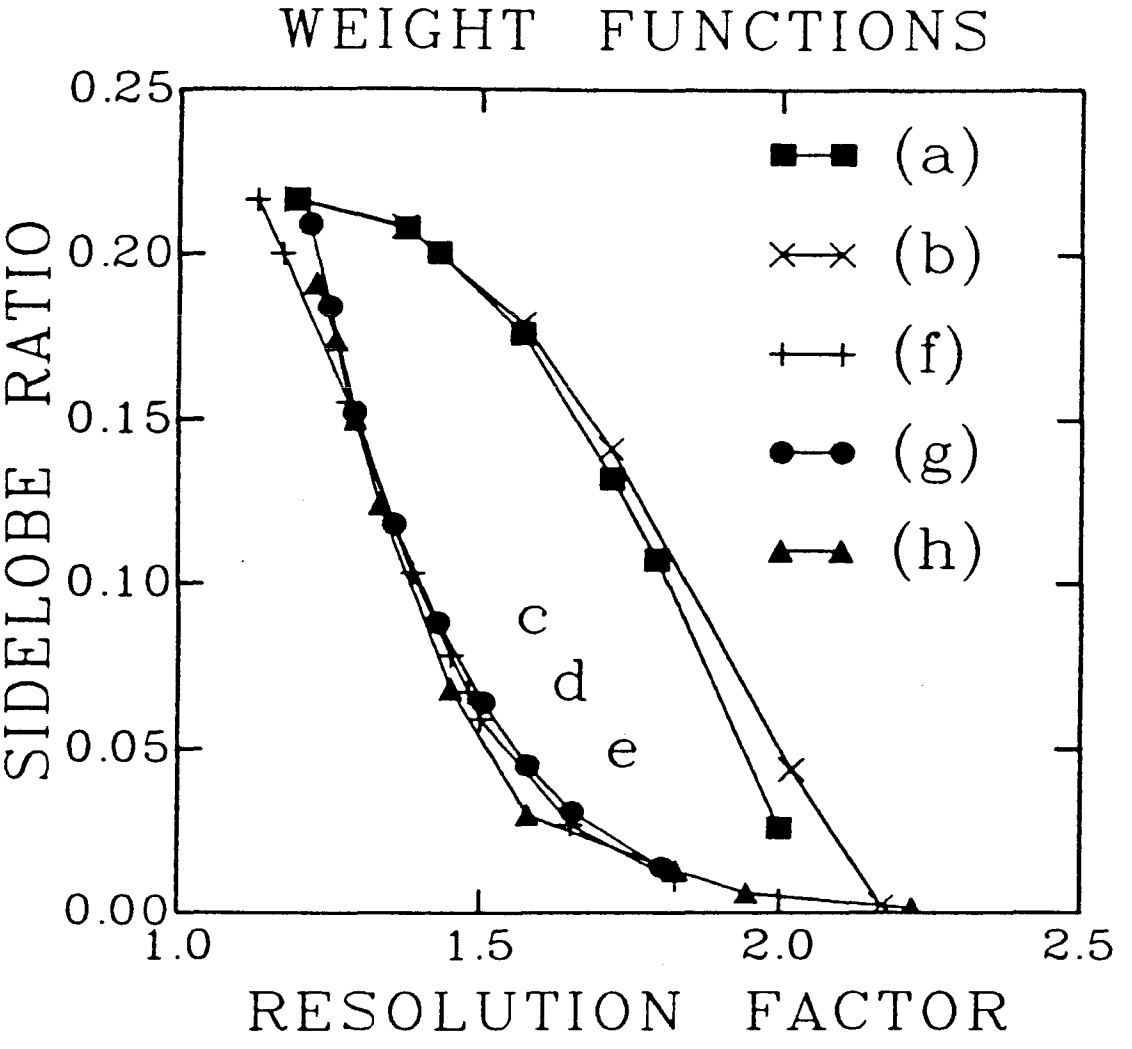
Figure 7. (a) Angle-resolved photoemission extended fine structure from S(1s) c(2x2)S/Ni(100) along [110]. The weighting function used for the taper is plotted offset above the data. Its minimum is zero and maximum is one. (b) Taper-and-transform Fourier amplitude for (a). (c) Autoregressive prediction of (a). An order M=64 was used for 128 points interpolated from the raw data. The weighting function is set above as for (a). (d) Fourier amplitude of the product of the window and extrapolated data in (c).

Figure 8. Selection of rank for the singular value decomposition for order 128 autoregressive fit to the data shown in figure 7. The singular value decomposition algorithm (ref. 17) automatically orders the singular values by size. The values σ_i^2, are plotted versus i as solid circles connected by a thin line; their magnitude is given by the left hand axis. The point by point differences are plotted as the thick line with the right hand axis giving the scale. The rank is selected as the point where the singular values depart from constant slope.



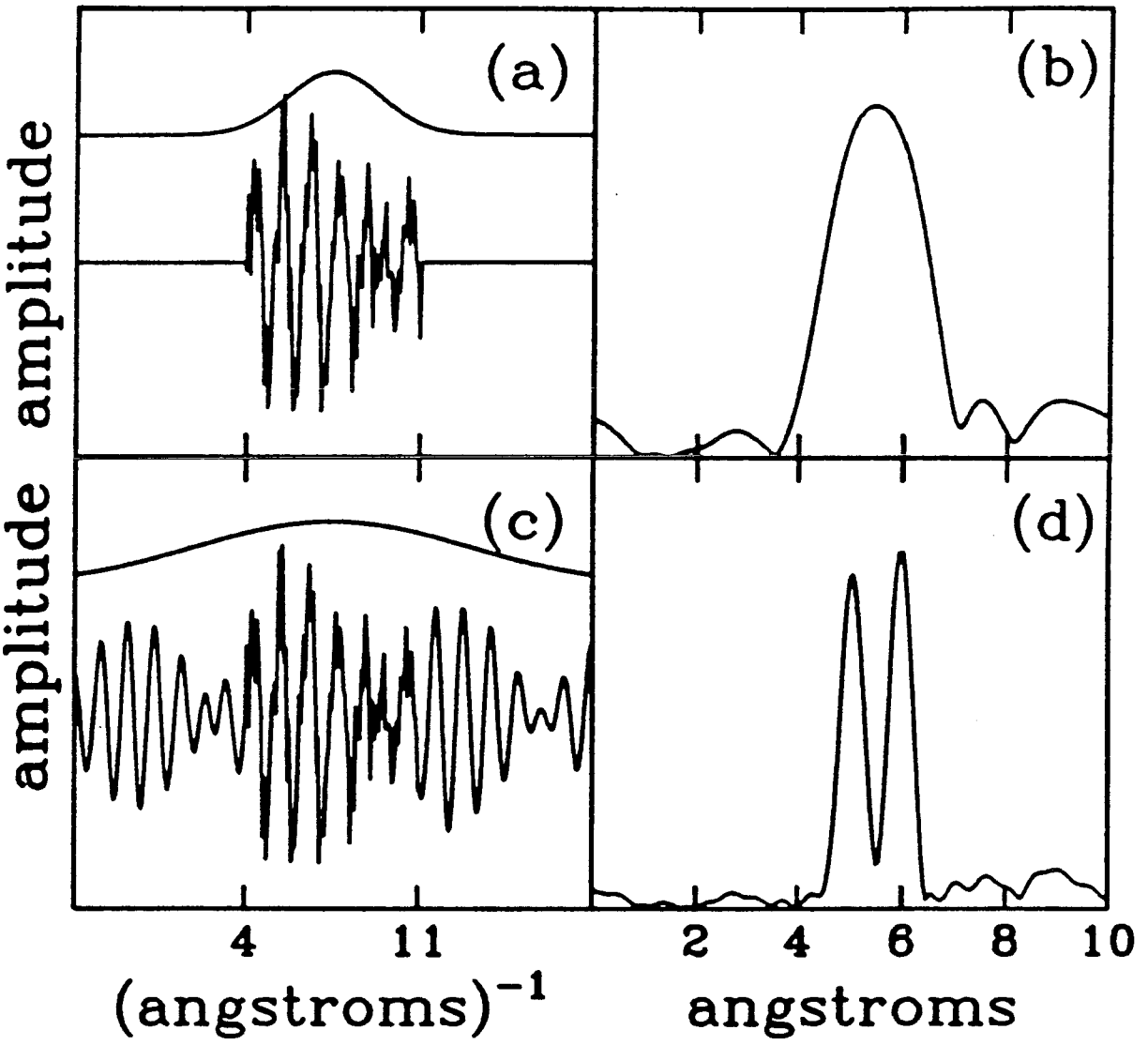
XBL 848-3552

Figure 1



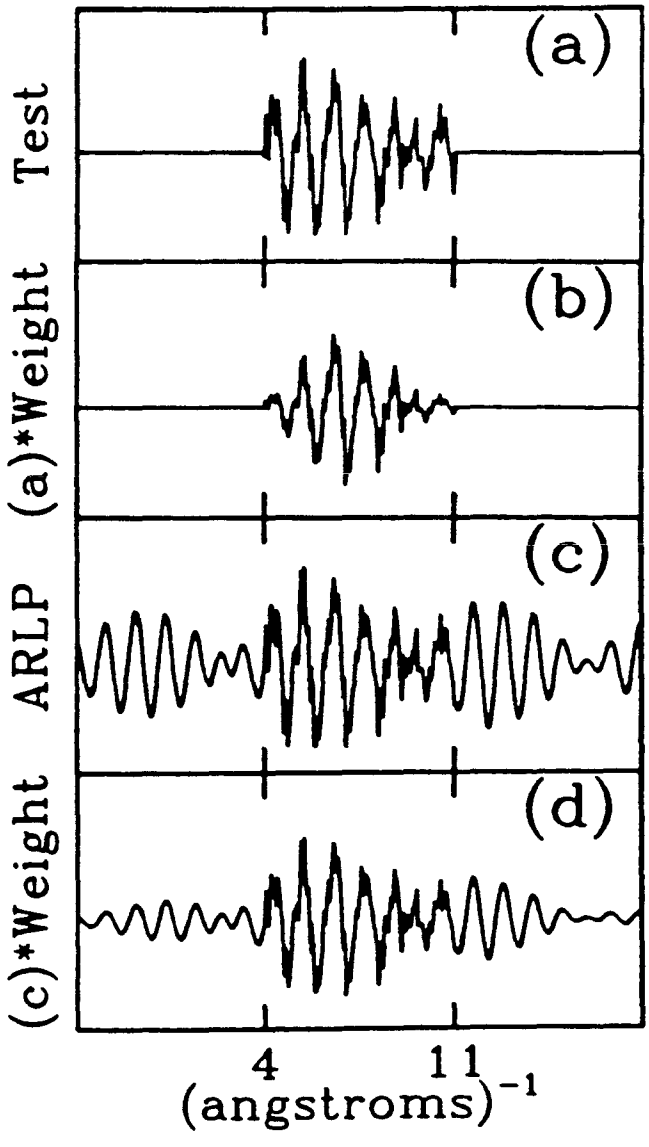
XBL 844-1349

Figure 2



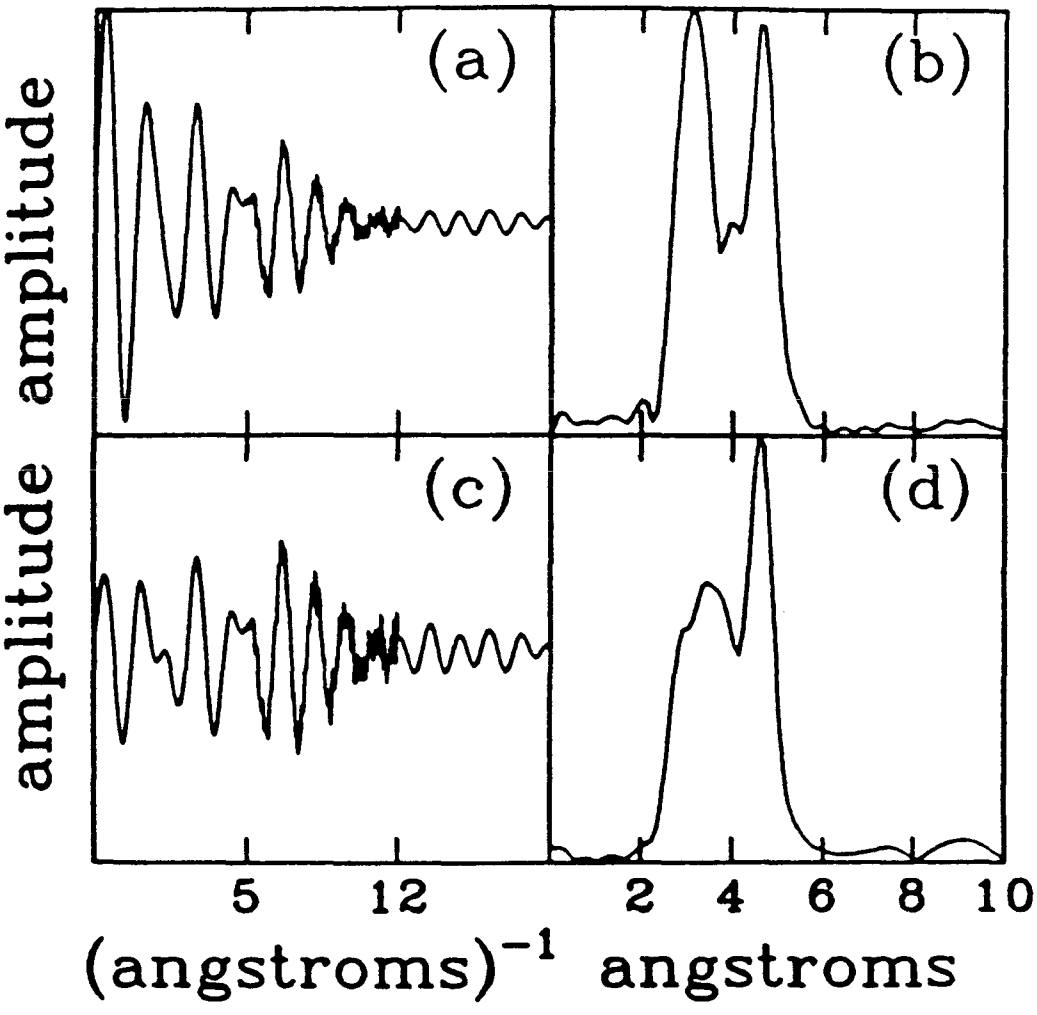
XBL 848-3553

Figure 3



XBL 848-3550

Figure 4



XBL 848-3551

Figure 5

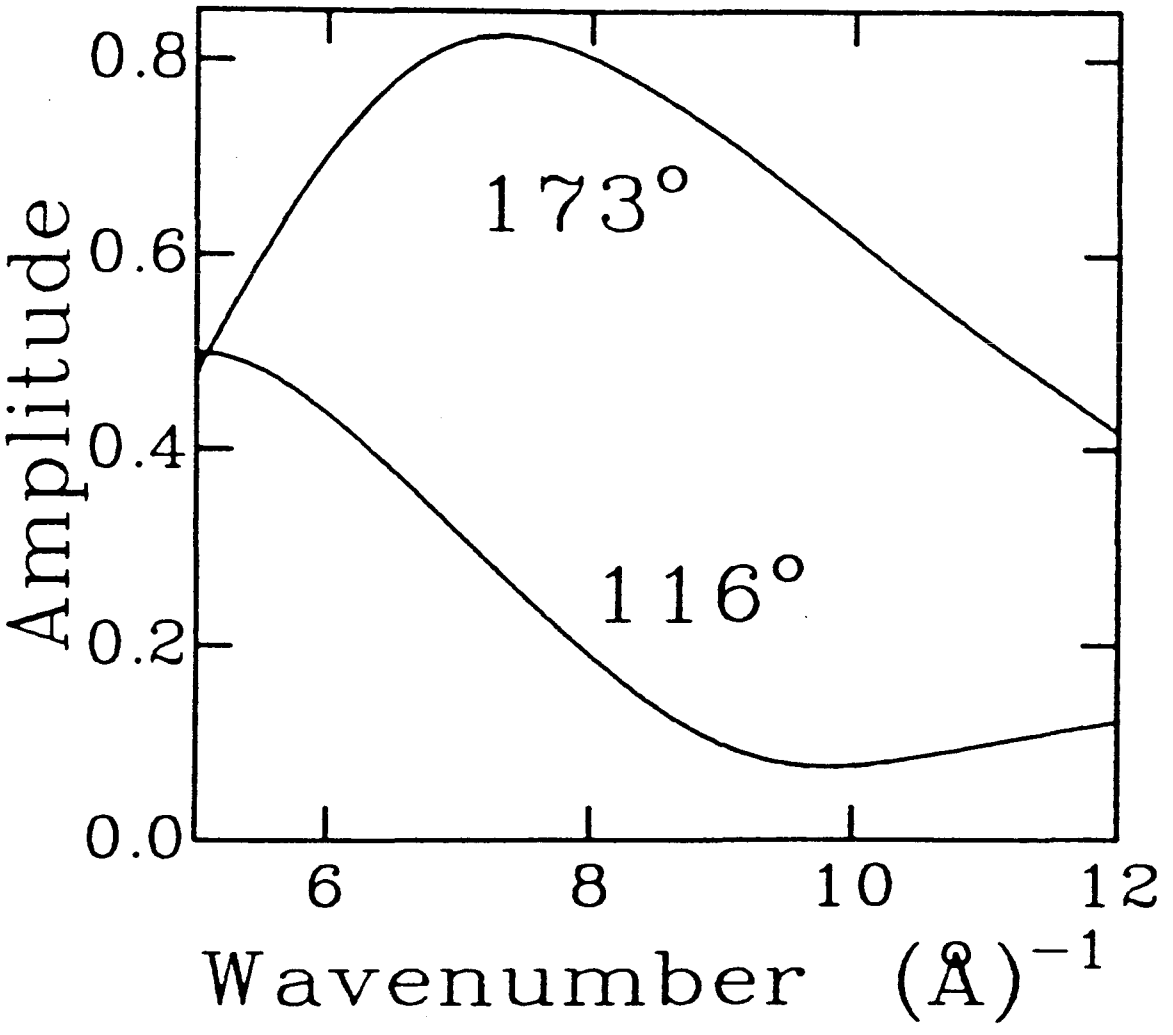
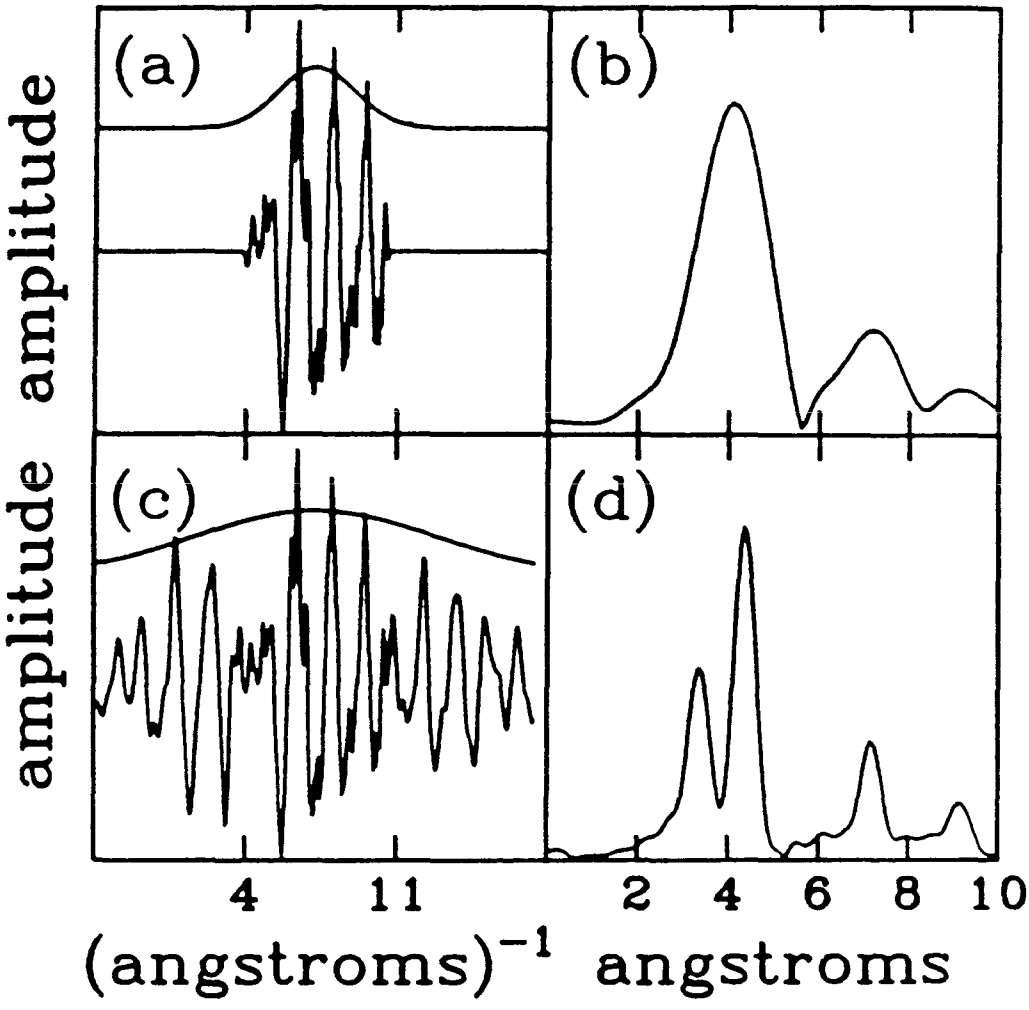


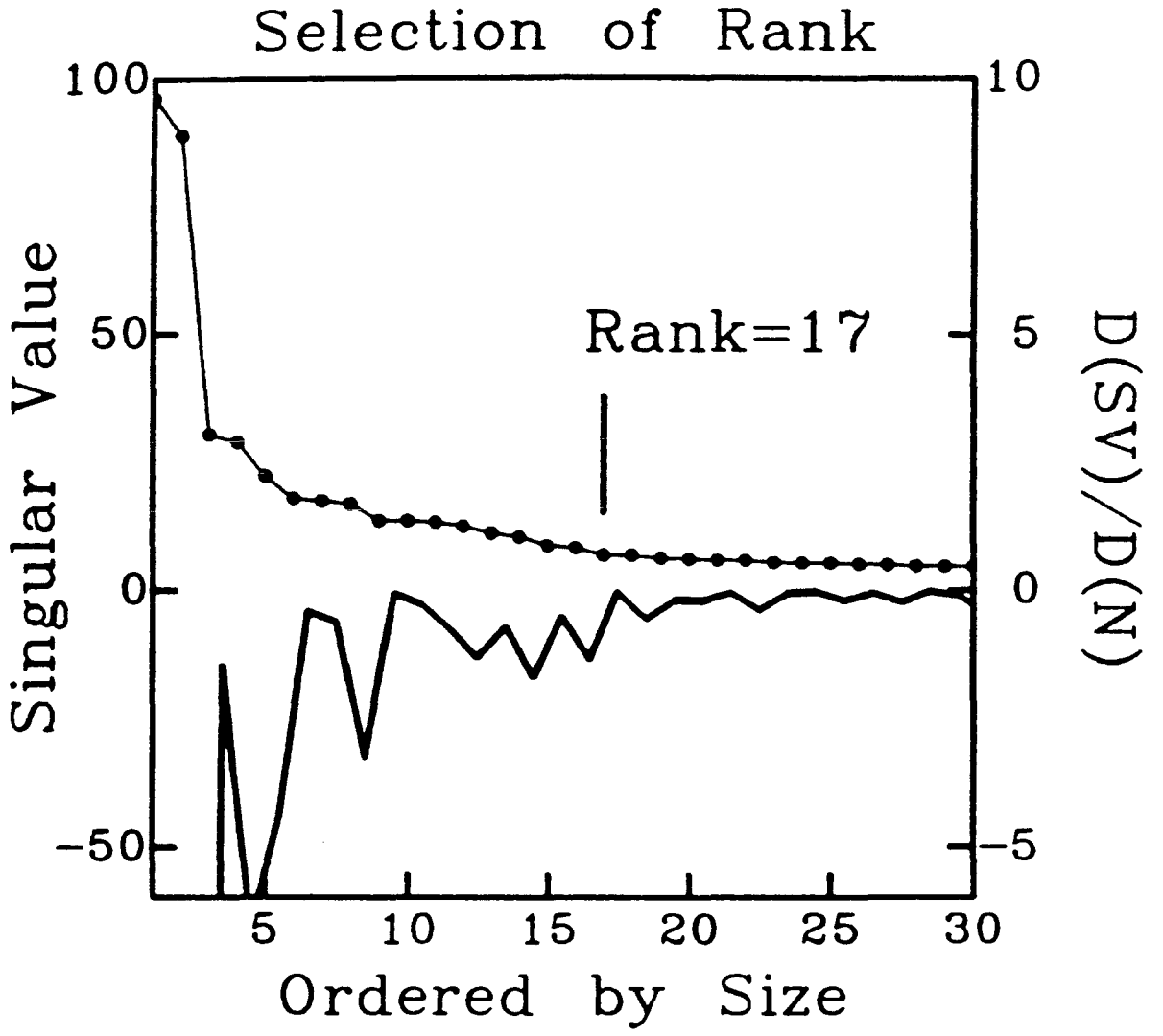
Figure 6

XBL 844-1350



XBL 848-3545

Figure 7



XBL 848-3549

Figure 8